On the Noise-Compensated Yule-Walker Equations

Carlos E. Davila

Electrical Engineering Department, Southern Methodist University
Dallas, Texas 75275-0338

E-mail: cd@seas.smu.edu
Phone: (214) 768-3197
Fax: (214) 768-3573

Abstract

Recently a method of estimating the parameters of an $AR(p)$ random process based on measurements corrupted by additive white noise was described. The method involves solving a matrix pencil, called the Noise-Compensated Yule-Walker (NCYW) equations, for the AR parameters and the variance of the measurement noise. In this correspondence we give a necessary and sufficient condition for there to exist a unique solution to the NCYW equations.

submitted to *IEEE Transactions on Signal Processing*
April 20, 2001

I. Background

The p^{th} -order AR ($AR(p)$) Random Process is given by

$$x(n) = -a(1)x(n-1) - a(2)x(n-2) - \cdots - a(p)x(n-p) + w(n) \quad (1)$$

where $w(n)$ is white noise having variance σ_w^2 and $a(k), k = 1, \dots, p$ are the AR parameters. We assume that $x(n)$ is real. The autocorrelation function of the AR process, $r_x(k)$, also satisfies the autoregressive property, this leads to the well-known *Yule-Walker* equations for the AR parameters

$$r_x(k) = -\sum_{i=1}^p a(i)r_x(k-i), \quad k \geq 1 \quad (2)$$

Suppose the measurements used to estimate the AR parameters can be modeled as $\tilde{x}(n) = x(n) + v(n)$ where $v(n)$ is white noise having variance σ_v^2 , then the parameter estimates derived from the Yule-Walker equations will be biased since, $r_{\tilde{x}}(k) = r_x(k) + \delta(k)\sigma_v^2$ where $\delta(k)$ is the Dirac delta function. It has been shown that the biased AR parameters produce a “flatter” AR spectrum since the estimated poles of the AR process are biased toward the center of the unit circle [1]. A number of methods have been described for estimating the AR parameters using noisy measurements, some of these methods are surveyed in [1, 5].

The Noise-Compensated Yule-Walker (NCYW) equations are defined as

$$(R_{\tilde{x}} - \lambda B)u = 0_{p+q} \quad (3)$$

where

$$R_{\tilde{x}} = \begin{bmatrix} r_{\tilde{x}}(1) & r_{\tilde{x}}(0) & r_{\tilde{x}}(-1) & \cdots & r_{\tilde{x}}(-(p-1)) \\ r_{\tilde{x}}(2) & r_{\tilde{x}}(1) & r_{\tilde{x}}(0) & \cdots & r_{\tilde{x}}(-(p-2)) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{\tilde{x}}(p) & r_{\tilde{x}}(p-1) & r_{\tilde{x}}(p-2) & \cdots & r_{\tilde{x}}(0) \\ r_{\tilde{x}}(p+1) & r_{\tilde{x}}(p) & r_{\tilde{x}}(p-1) & \cdots & r_{\tilde{x}}(1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{\tilde{x}}(p+q) & r_{\tilde{x}}(p+q-1) & r_{\tilde{x}}(p+q-2) & \cdots & r_{\tilde{x}}(q) \end{bmatrix} \quad (4)$$

$$B = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (5)$$

The dimensions of $R_{\hat{x}}$, B , and u are $(p+q) \times (p+1)$, $(p+q) \times (p+1)$, and $(p+1) \times 1$, respectively. The unknowns in (3) are u and λ , so clearly, equality holds when $\lambda = \sigma_v^2$ and $u = \begin{bmatrix} 1 & a(1) & a(2) & \cdots & a(p) \end{bmatrix}^T$. We observe that the first p equations are nonlinear in the AR parameters, u , and the measurement noise variance, λ . The next q equations however are linear in u . In general there exist p distinct (λ, u) solving (3), the solution is taken to be that which corresponds to the minimum real value of λ solving (3).

A number of authors have observed that solving the p nonlinear equations, in addition to the q linear equations can improve the parameter estimates [2, 3, 4]. In [5], a matrix pencil method based on the Noise-Compensated Yule-Walker (NCYW) equations was presented which was found to out-perform several other methods for *AR*-in noise parameter estimation. None of these papers have established the minimum number of linear equations that are required for the solution of the NCYW equations to be the unique, correct solution. It is clear that $q \geq p$ linear equations are sufficient to insure that the solution is unique, in this case, all other (λ, u) solving (3) are complex. However, the minimum value of q needed to insure a unique solution has not been established. This is an important consideration because using a large value of q , which would guarantee a unique solution, also implies computing more high-order autocorrelation lag estimates which can reduce the solution accuracy since these tend to have a larger estimation variance. Hence one would like to choose the smallest value of q that still guarantees a unique solution. One might expect that since there are a total of $p+q$ equations in $p+1$ unknowns, fewer than $q=p$ linear equations are needed. In other words, if only one linear equation were used, $q=1$, this would still give $p+1$ equations in $p+1$ unknowns, hence a unique solution could still exist. In this correspondence, we show that this is not the case and $q \geq p$ is also a necessary condition for there to exist a unique

solution.

II. $AR(p)$ Uniqueness

We wish to determine whether there exist two *distinct* $AR(p)$ random processes, $x(n)$ and $y(n)$, having autocorrelation functions $r_x(k)$ and $r_y(k)$, for which

$$\begin{aligned} r_x(0) &\neq r_y(0) \\ r_x(k) &= r_y(k), k = 1, 2, \dots, p, p+1, \dots, 2p-1 \end{aligned} \quad (6)$$

By distinct processes, we mean that the polynomials $1 + \sum_{i=1}^p a(i)z^{-i}$ for the two $AR(p)$ processes differ by at least one root. Clearly, if two such processes exist, then we cannot find a unique solution to the NCYW equations with $q \leq p-1$ linear equations. For example, suppose that $p=3$, $q=p-1=2$, $r_x(k) = \{1, r(1), r(2), r(3), r(4), r(5)\}$, and $r_y(k) = \{1.5, r(1), r(2), r(3), r(4), r(5)\}$. Now let the observed noisy process have autocorrelation function $r_{\tilde{x}}(k) = \{2, r(1), r(2), r(3), r(4), r(5)\}$. Then, for $q=p-1$, the NCYW equations would yield two valid solutions, one corresponding to the AR parameters of $x(n)$, and $\lambda=1$, and the other corresponding to the AR parameters of $y(n)$ and $\lambda=0.5$. We do not make any assumptions about the nature of the $AR(p)$ process whose parameters we are trying to estimate. We simply seek the minimum value of q for which the NCYW equations will produce a unique solution for *any* $AR(p)$ random process. Hence we need only find one example of two distinct $AR(p)$ processes satisfying (6) to establish that the minimum value of q giving a unique solution is p .

The two autocorrelation functions $r_x(k)$ and $r_y(k)$ have the z -domain representation

$$R_x(z) = \frac{\sigma_w^2}{(1 - \gamma_1 z^{-1})(1 - \gamma_2 z^{-1}) \cdots (1 - \gamma_p z^{-1})(1 - \gamma_1^* z)(1 - \gamma_2^* z) \cdots (1 - \gamma_p^* z)} \quad (7)$$

and

$$R_y(z) = \frac{\sigma_z^2}{(1 - \gamma_{p+1} z^{-1})(1 - \gamma_{p+2} z^{-1}) \cdots (1 - \gamma_{2p} z^{-1})(1 - \gamma_{p+1}^* z)(1 - \gamma_{p+2}^* z) \cdots (1 - \gamma_{2p}^* z)} \quad (8)$$

respectively, where σ_w^2 and σ_z^2 is the power of the white noise in the $AR(p)$ model. We can

parameterize $x(n)$ and $y(n)$ by, $(\sigma_w^2, \gamma_1, \dots, \gamma_p)$ and $(\sigma_z^2, \gamma_{p+1}, \dots, \gamma_{2p})$, respectively. If we assume that all $2p$ poles, $\gamma_1, \dots, \gamma_{2p}$ are distinct, there must exist a vector β for which

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ \gamma_1 & \gamma_2 & \cdots & \gamma_p & \gamma_{p+1} & \gamma_{p+2} & \cdots & \gamma_{2p} \\ \gamma_1^2 & \gamma_2^2 & \cdots & \gamma_p^2 & \gamma_{p+1}^2 & \gamma_{p+2}^2 & \cdots & \gamma_{2p}^2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_1^{2p-1} & \gamma_2^{2p-1} & \cdots & \gamma_p^{2p-1} & \gamma_{p+1}^{2p-1} & \gamma_{p+2}^{2p-1} & \cdots & \gamma_{2p}^{2p-1} \end{bmatrix} \beta = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (9)$$

since the Vandermonde matrix in (9) is non-singular [6]. The vector $\beta = [\beta_1 \ \beta_2 \ \cdots \beta_{2p}]^T$ is the first column of the inverse of the Vandermonde matrix in (9) and is given by [7]

$$\beta_i = \frac{-\prod_{\substack{j=1 \\ j \neq i}}^{2p} \gamma_j}{\prod_{\substack{j=1 \\ j \neq i}}^{2p} (\gamma_i - \gamma_j)}, i = 1, \dots, 2p \quad (10)$$

The autocorrelation functions $r_x(k)$ and $r_y(k)$ can be expressed as

$$r_{x(y)}(k) = \frac{1}{j2\pi} \oint_{\mathcal{C}} R_{x(y)}(z) z^{k-1} dz \quad (11)$$

where the contour of integration $\mathcal{C}$ is the unit circle. Using the Cauchy Integral Formula [8],

$$r_{x(y)}(k) = \sum \left[\text{residues of } R_{x(y)}(z) z^{k-1} \text{ at } \gamma_{i(p+i)}, i = 1, \dots, p \right] \quad (12)$$

The residues of $R_x(z) z^{k-1}$ at $\gamma_i, i = 1, \dots, p$ are given by [8]

$$\alpha_{i,k} = \frac{\sigma_w^2 \gamma_i^{k-1}}{\prod_{\substack{j=1 \\ j \neq i}}^p (\gamma_i - \gamma_j) \prod_{k=1}^p (\gamma_i^{-1} - \gamma_k^*)}, i = 1, \dots, p \quad (13)$$

while the residues of $R_y(z) z^{k-1}$ at $\gamma_i, i = p+1, \dots, 2p$ are

$$\alpha_{i,k} = \frac{\sigma_z^2 \gamma_i^{k-1}}{\prod_{\substack{j=p+1 \\ j \neq i}}^{2p} (\gamma_i - \gamma_j) \prod_{k=p+1}^{2p} (\gamma_i^{-1} - \gamma_k^*)}, i = p+1, \dots, 2p \quad (14)$$

Now if (6) is satisfied, using (12)-(14) leads to

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ \gamma_1 & \gamma_2 & \cdots & \gamma_p & \gamma_{p+1} & \gamma_{p+2} & \cdots & \gamma_{2p} \\ \gamma_1^2 & \gamma_2^2 & \cdots & \gamma_p^2 & \gamma_{p+1}^2 & \gamma_{p+2}^2 & \cdots & \gamma_{2p}^2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_1^{2p-1} & \gamma_2^{2p-1} & \cdots & \gamma_p^{2p-1} & \gamma_{p+1}^{2p-1} & \gamma_{p+2}^{2p-1} & \cdots & \gamma_{2p}^{2p-1} \end{bmatrix} \alpha = \begin{bmatrix} K \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (15)$$

where K is a constant and $\alpha = \begin{bmatrix} \alpha_{1,0} & \cdots & \alpha_{p,0} & -\alpha_{p+1,0} & \cdots & -\alpha_{2p,0} \end{bmatrix}^T$. Hence if there exist distinct poles $\gamma_i, i = 1, 2, \dots, 2p$, and σ_w^2 and σ_z^2 , such that $\alpha = K\beta$, then the NCYW equations do not possess a unique solution for $q = p - 1$. Equating the vectors α and $K\beta$ leads to

$$\frac{\sigma_w^2 K^{-1} \prod_{j=p+1}^{2p} (\gamma_i - \gamma_j)}{\prod_{j=1}^p (\gamma_i^{-1} - \gamma_j^*)} = - \prod_{j=1}^{2p} \gamma_j, i = 1, \dots, p \quad (16)$$

and

$$\frac{\sigma_z^2 K^{-1} \prod_{j=1}^p (\gamma_i - \gamma_j)}{\prod_{j=p+1}^{2p} (\gamma_i^{-1} - \gamma_j^*)} = \prod_{j=1}^{2p} \gamma_j, i = p + 1, \dots, 2p \quad (17)$$

It can be readily verified that these equations are satisfied for any value of p by choosing

$$\gamma_k = a e^{j2\pi(k-1)/p}, k = 1, \dots, p \quad (18)$$

and

$$\gamma_k = b e^{j2\pi(k-p-1)/p}, k = p + 1, \dots, 2p \quad (19)$$

where $0 < b < a < 1$. Hence, since two possible models exist which satisfy (6) for $q = p - 1$ (and hence for all smaller q), then a necessary condition for the NCYW equations to have a unique solution is that $q \geq p$.

III. Numerical Example

Consider the case where $p = 3$, $a = 0.9$, and $b = 0.8$. The six poles are, according to (18) and (19), $\gamma_1 = 0.9, \gamma_2 = 0.9e^{j2\pi/3}, \gamma_3 = 0.9e^{-j2\pi/3}, \gamma_4 = 0.8, \gamma_5 = 0.8e^{j2\pi/3}$, and $\gamma_6 = 0.8e^{-j2\pi/3}$.

Equations (16) and (17) are satisfied with this choice of poles if we choose, $\sigma_w^2 = 1.105540129$, $\sigma_z^2 = 2.478788129$, and $K = 1$. The autocorrelation function ($k = 0, \dots, 5$) for these two $AR(3)$ models are found in Table 1. If, for example the noisy $AR(3)$ random process, $\tilde{x}(n)$, has $r_{\tilde{x}}(0) = 4.3594470046$, and all other lags equal to those in Table 1, then for $q = p - 1 = 2$, there would be two valid solutions to the NCYW equations, one corresponding to a additive noise variance of $\sigma_v^2 = 1$, and the other corresponding to $\sigma_v^2 = 2$.

k	$r_x(k)$	$r_y(k)$
0	2.3594470046	3.3594470046
1	0	0
2	0	0
3	1.7200368664	1.7200368664
4	0	0
5	0	0

Table 1: Results of numerical example.

IV. Conclusions

It was shown that a necessary condition for the Noise-Compensated Yule-Walker (NCYW) equations to possess a unique solution is for there to be at least $q = p$ linear equations. Using fewer than p linear equations may lead to a non-unique solution. This does not mean that the p nonlinear equations should be dropped altogether. As has been demonstrated previously [2, 3, 4, 5], these nonlinear equations are composed of autocorrelation lag estimates which tend to be more accurate than those making up the linear equations and can be exploited to improve the accuracy of the autoregressive parameter estimates.

References

- [1] S. M. Kay, “Noise compensation for autoregressive spectral estimates,” *IEEE Trans. on Acoustics, Speech, and Signal Proc.*, vol. ASSP-28, pp. 292–303, June 1980.
- [2] V. K. Jain and B. S. Atal, “Robust LPC analysis of speech by extended correlation matching,” in *Proc. ICASSP 85*, pp. 473–476, March 1985.

- [3] K. K. Paliwal, “A noise-compensated long correlation matching method for ar spectral estimation of noisy signals,” in *Proc. ICASSP 86*, pp. 1369–1372, April 1986.
- [4] L. B. Jackson, J. Huang, K. P. Richards, and H. Chen, “AR, ARMA and AR-in-noise modeling by fitting windowed correlation data,” *IEEE Trans. Acoust. , Speech, Signal Processing*, vol. 37, pp. 1608–1612, Oct. 1989.
- [5] C. E. Davila, “A subspace approach to estimation of autoregressive parameters from noisy measurements,” *IEEE Trans. Signal Processing*, vol. SP-46, pp. 531–534, February 1998.
- [6] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Johns Hopkins University Press, 1989.
- [7] F. A. Graybill, *Matrices with Applications in Statistics*. Wadsworth, 1983.
- [8] R. V. Churchill and J. W. Brown, *Complex Variables and Applications*. McGraw Hill, 1984.